Comp. & Maths. with Appls. Vol. 7, pp. 43–58
Pergamon Press Ltd., 1981. Printed in Great Britain

DUALITY METHODS FOR SOLVING VARIATIONAL INEQUALITIES†

A. BERMÚDEZ and C. MORENO
Departamento de Ecuaciones Funcionales, Universidad de Santiago, Spain

Communicated by R. Glowinski

(*Received December* 1979)

Abstract—Methods of maximal monotone operators are used in order to study, from a general point of view, duality numerical algorithms for solving variational inequalities. With classical algorithms, such as Uzawa's method for the standard and augmented Lagrangian, this paper presents some new algorithms, which appear to have very good numerical performances.

1. INTRODUCTION

Let V be a real Hilbert space and A a monotone operator, $A \in L(V, V')$. Let ψ be a lower semi-continuous convex function over V which takes its values in $(-\infty, \infty]$ and which is not identically $+\infty$.

If A is symmetric, every solution of the variational inequality (where $f \in V'$):

$$\begin{aligned}&(Ay, z-y) + \psi(z) - \psi(y) \geq (f, z-y) \quad \text{for all } z \in V\\ &y \in V\end{aligned} \tag{1.1}$$

achieves the infimum of the functional:

$$z \in V \to J(z) = \frac{1}{2}(Az, z) - (f, z) + \psi(z). \tag{1.2}$$

Duality methods for the numerical solution of (1.1) try to overcome the difficulty related to the non differentiability of ψ by constructing a convenient lagrangian.

Let ψ^* be the *conjugate functional* of ψ, i.e.

$$\psi^*(z) = \sup_{q \in V} \{(z, q) - \psi(q)\},$$

then for any $z \in V$ we have:

$$J(z) = \frac{1}{2}(Az, z) - (f, z) + \sup_{q \in V} \{(q, z) - \psi^*(q)\}. \tag{1.3}$$

Hence:

$$\inf_{z \in V} J(z) = \inf_{z \in V} \sup_{q \in V} L(z, q), \tag{1.4}$$

where L is the Lagrangian defined by

$$L(z, q) = \frac{1}{2}(Az, z) - (f, z) + (q, z) - \psi^*(q). \tag{1.5}$$

†This research was carried at I.R.I.A. (France) during a visit sponsored by Scholarship of the French Embassy in Spain.

If $\{y, p\}$ is a saddle-point of L, then y is a solution of (1.1) and moreover:

$$\begin{aligned} &Ay + p = f \\ &p \in \partial\psi(y), \end{aligned} \tag{1.6}$$

where $\partial\psi$ is the *subdifferential* of ψ.

In several examples, ψ is the *support function* of a closed convex set K, i.e.

$$\psi(z) = \sup_{q \in K} \{(z, q)\}.$$

In these cases, ψ^* is the *indicator function* of K and we may use the Uzawa's algorithm to compute y:

$$\begin{aligned} &Ay^m + p^m = f \\ &p^{m+1} = P_K(p^m + \rho_m y^m). \end{aligned} \tag{1.7}$$

From now on A is not necessarily symmetric. If y is a solution of (1.1) there exists $p \in \partial\psi(y)$ such that:

$$Ay + p = f. \tag{1.8}$$

Besides, we have (see Bermudez–Moreno[1]):

$$p \in \partial\psi(y) \textit{ if an only if } p = G_\lambda(y + \lambda p) \textit{ for every } \lambda > 0, \tag{1.9}$$

where G_λ is the *Yosida approximation* of the maximal monotone operator $\partial\psi$ (see Pazy[2]). Thus, it seems appropriate to define the following algorithm:

Let p^0 be given arbitrarily. Given p^m, we define y^m, p^{m+1} by:

$$\begin{aligned} &Ay^m + p^m = f \\ &p^{m+1} = G_{\lambda_m}(y^m + \lambda_m p^m). \end{aligned} \tag{1.10}$$

It is easy to see that (1.10) is in fact (1.7) if ψ is the support function of the closed convex set K.

Observe that (1.10) may always be defined since neither the G-differentiability of the functional ψ^*, nor the alternative condition "ψ^* is the indicator function of K", nor even the symmetry of A are necessary. Thus (1.9) provids an elegant way of getting some numerically efficient algorithms.

Under quite general hypotheses we prove, in Section 3, the convergence of the algorithm (1.10), which is presented as a particular case of a certain class of algorithms which are obtained by perturbing the operator $\partial\psi$.

From (1.9) and the formulation (1.6), we also obtain, in Section 3, the penalty-duality algorithm, i.e. the Uzawa's method for the augmented Lagrangian (see Hestenes[3], Powell[4], Glowinski–Marrocco[5], Fortin[6]). The practical use of this algorithm is shown there and we give several variants easily implementable on computers.

We place our report in a slightly different context to the one defined at the beginning of this introduction. Actually, taking E a Hilbert space, φ a lower semi-continuous proper convex function in E, B a bounded linear operator from E into V' and Λ_E the canonical isomorphism from E into E', we study the more general case where $\psi = \varphi \circ \Lambda_E^{-1} \circ B^*$. This modification is very interesting for certain applications, as is shown in the paper.

2. THEORETICAL BACKGROUND

2.1 $M(\omega)$ operators (*see Pazy*[2])

Let E be a real Hilbert space and ω a real number. A multivalued operator G in E is called a *maximal-$M(\omega)$ operator* if $G + \omega I$ is maximal monotone.

If $\lambda\omega < 1$, then the operator $J_\lambda = (I + \lambda G)^{-1}$ is defined over all E and univalued. Moreover it is a monotone Lipschitz function with constant $(1 - \lambda\omega)^{-1}$.

As in the case of maximal monotone operators we may define the Yosida approximation of G by:

$$G_\lambda = \frac{I - J_\lambda}{\lambda}. \tag{2.1}$$

It is not difficult to see that $G_\lambda \in M(\omega/1 - \lambda\omega)$ and we also have:

$$\frac{1 - 2\lambda\omega}{\lambda^2}\|J_\lambda(v_1) - J_\lambda(v_2)\|^2 + \|G_\lambda(v_1) - G_\lambda(v_2)\|^2 \leqslant \frac{1}{\lambda^2}\|v_1 - v_2\|^2 \tag{2.2}$$

and

$$\|G_\lambda(v_1) - G_\lambda(v_2)\|^2 \leqslant \frac{1}{\lambda}(G_\lambda(v_1) - G_\lambda(v_2), v_1 - v_2)$$
$$+ \frac{\omega}{\lambda}\|J_\lambda(v_1) - J_\lambda(v_2)\|^2 \tag{2.3}$$

where v_1, $v_2 \in E$.

In particular, if $\lambda\omega \leqslant (1/2)$, G_λ is a Lipschitz function with constant $1/\lambda$.

The following lemmas are fundamental throughout this paper. The first one has its interest in constructing an equivalent formulation for the problem (1.6). The second lemma establishes an inequality of great interest to prove the theorems of convergence.

LEMMA 2.1

Let G a maximal $M(\omega)$-operator. If $\lambda\omega < 1$, the following statements are equivalent

(i) $u \in G(v)$

(ii) $u = G_\lambda(v + \lambda u)$.

Proof

See Bermudez–Moreno[1].

LEMMA 2.2

Under the hypotheses of Lemma 2.1, and for v_1, v_2, w_1, w_2 in E, we have:

$$\left\|v_1 - v_2 - \frac{1}{\lambda}(J_\lambda(w_1) - J_\lambda(w_2))\right\|^2 \leqslant \frac{1}{\lambda^2}\|w_1 - w_2 - \lambda(v_1 - v_2)\|^2$$
$$+ \|v_1 - v_2\|^2 + \frac{2\omega}{\lambda}\|J_\lambda(w_1) - J_\lambda(w_2)\|^2. \tag{2.4}$$

Proof

From (2.1) we have:

$$\frac{1}{\lambda^2}\|\lambda(v_1 - v_2) - (J_\lambda(w_1) - J_\lambda(w_2))\|^2 = \frac{1}{\lambda^2}\|\lambda(v_1 - v_2) - (w_1 - w_2)\|^2$$
$$+ \|G_\lambda(w_1) - G_\lambda(w_2)\|^2 + \frac{2}{\lambda}(\lambda(v_1 - v_2) - (w_1 - w_2), G_\lambda(w_1) - G_\lambda(w_2)). \tag{2.5}$$

Then the lemma follows taking account of this equality and (2.3).

2.2 *Elliptic variational inequalities*

Let V be a real Hilbert space and $A \in L(V, V')$ such that:

$$(Az, z)_{V'V} \geq \alpha \|z\|_V^2 \quad \text{for all } z \in V, \tag{2.6}$$

where $\alpha \geq 0$.

Let us consider now an operator $B \in L(E, V')$ and φ a proper lower semi-continuous (l.s.c.) convex function over E (see Ekeland–Temam [7]). We assume that

$$\varphi((1-\lambda)v_1 + \lambda v_2) \leq (1-\lambda)\varphi(v_1) + \lambda\varphi(v_2) - \frac{1}{2}\beta\lambda(1-\lambda)\|v_1 - v_2\|^2, \tag{2.7}$$

for all $v_1, v_2 \in E$ and $\lambda \in (0, 1)$, where $\beta \geq 0$ (if $\beta > 0$, φ is called a strongly convex function with modulus β).

Finally, let ψ be the l.s.c. convex function

$$\psi = \varphi \circ \Lambda_E^{-1} \circ B^* \tag{2.8}$$

which we suppose proper.

Throughout this paper, it is assumed that, at least, one of the following hypotheses holds:

(H1) $\alpha > 0$.

(H2) $\beta > 0$ and $B\Lambda_E^{-1}B^*$ is an isomorphism from V onto V'.

With (H1) or (H2), the variational inequality:

$$(Ay, z - y)_{V'V} + \psi(z) - \psi(y) \geq (f, z - y)_{V'V} \quad \text{for all } z \in V \tag{2.9}$$

has a unique solution $y \in V$.

Related to this variational inequality we consider the following problem:

$$\text{To find } y \in V \text{ such that} \quad f - Ay \in B\partial\varphi(\Lambda_E^{-1}B^*y). \tag{2.10}$$

Since the inclusion:

$$B(\partial\varphi(\Lambda_E^{-1}B^*z)) \subset \partial\psi(z) \tag{2.11}$$

is always satisfied for all $z \in V$, if y is a solution of (2.10) then y is the solution of (2.9). On the contrary, the reciprocal inclusion is not true in the general case, (see Ekeland–Temam [7] for sufficient conditions). Because of this, the problems (2.9) and (2.10) are not always equivalent. However, we have that equivalence for almost all the classical examples, at least if they are considered in their discrete form.

From now on, we are going to suppose that the problem (2.10) has solution. For the above reasons, this solution is unique and agrees with the solution of (2.9).

2.3 *Examples*

Let Ω a bounded open set of R^n with "smooth" boundary Γ.

2.3.1 *Flow of a Bingham fluid in a cylindrical pipe.* (Duvaut–Lions [8, Chap. 6], Glowinski–Lions–Tremolieres [9, Chap. 5]).

It consists of finding $y \in H_0^1(\Omega)$ which minimizes the functional

$$J(z) = \frac{\nu}{2}\int_\Omega |\text{grad } z|^2 \,dx + \int_\Omega |\text{grad } z| \,dx - \int_\Omega fz \,dx, \tag{2.12}$$

where $\nu > 0$ and $f \in L^2(\Omega)$.

It is easy to check that y is also the solution of the problem (2.10) with $V = H_0{}^1(\Omega)$, $E = (L^2(\Omega))^n$,

$$(Ay, z) = \nu \int_\Omega \sum_{i=1}^{n} \frac{\partial y}{\partial x_i} \frac{\partial z}{\partial x_i} \, dx. \tag{2.13}$$

$B = -\text{div}$, and φ being the support function of the closed convex set:

$$K = \{v \in (L^2(\Omega))^n : |v(x)| \leq 1 \quad \text{a.e. in } \Omega\}. \tag{2.14}$$

2.3.2 *Elastoplastic torsion* (Glowinski–Lions–Tremolieres [9, Chap. 3]. In this problem we wish to find a y which achieves the minimum of the functional:

$$J(z) = \frac{\nu}{2} \int_\Omega |\text{grad } z|^2 \, dx - \int_\Omega fz \, dx \tag{2.15}$$

over the closed convex set:

$$C = \{z \in H_0{}^1(\Omega) : |\text{grad } z| \leq 1 \quad \text{a.e. in } \Omega\}.$$

If f is a constant function, then y is the solution of the problem (2.10) with V, E, A, B, K as in 2.3.1, and φ the indicator function of the set K in (2.14) (see Brezis [10]).

3. NUMERICAL ALGORITHMS

3.1 *A linear iterative algorithm*

Let ω be an arbitrary real number. We denote by G^ω the operator $\partial\varphi - \omega I$. Clearly G^ω is a maximal-$M(\omega - \beta)$ operator; therefore, if $\lambda(\omega - \beta) < 1$, we may define the resolvent operator $J_\lambda{}^\omega$ and the Yosida approximation $G_\lambda{}^\omega$. Besides it is easy to prove that:

$$G_\lambda{}^\omega(v) = \frac{1}{1 - \lambda\omega} G_{\lambda/(1-\lambda\omega)}\left(\frac{1}{1 - \lambda\omega} v\right) - \frac{\omega}{1 - \lambda\omega} v. \tag{3.1}$$

On the other hand, by the Lemma 2.1, the problem (2.10) is equivalent to
Finding $y \in V$ such that:

$$Ay + \omega B \Lambda_E{}^{-1} B^* y = f - Bu \tag{3.2}$$

$$u = G_\lambda{}^\omega(\Lambda_E{}^{-1} B^* y + \lambda u). \tag{3.3}$$

Remark 3.1. In general, if y is the solution of the problem (2.7), there exists more than one $u \in E$ satisfying (3.2) and (3.3). Clearly, sufficient conditions for the uniqueness are:
B is one-to-one in $\partial\varphi(\Lambda_E{}^{-1} B^* y)$ or φ is G-differentiable in $\Lambda_E{}^{-1} B^* y$.
The formulation (3.2), (3.3) suggests the following:

Algorithm 1
Let u^0 be arbitrary in E.

$$Ay^m + \omega B \Lambda_E{}^{-1} B^* y^m = f - Bu^m \tag{3.4}$$

$$u^{m+1} = G_{\lambda_m}^\omega(\Lambda_E{}^{-1} B^* y^m + \lambda_m u^m). \tag{3.5}$$

We state now a convergence result for this algorithm.

PROPOSITION 3.1
Under the hypotheses:

$$\lambda_m(\omega - \beta) \leq \frac{1}{2} \tag{3.6}$$

and

$$0<\epsilon_1\leqslant\frac{1}{\lambda_m}\leqslant\epsilon_2<2\left(\omega+\frac{\alpha}{\|B^*\|^2}\right) \tag{3.7}$$

we have:

$$\lim_{m\to\infty}\{y^m\}=y \tag{3.8}$$

where y is the solution of the problem (2.10).

Proof

From the inequality (2.2) we have:

$$\begin{aligned}&\frac{1-2\lambda_m(\omega-\beta)}{\lambda_m^2}\|J^{\omega}_{\lambda_m}(\Lambda_E^{-1}B^*y+\lambda_m u)-J^{\omega}_{\lambda_m}(\Lambda_E^{-1}B^*y^m+\lambda_m u^m)\|_E^2\\&+\|u-u^{m+1}\|_E^2\leqslant\frac{1}{\lambda_m^2}\|\Lambda_E^{-1}B^*(y-y^m)+\lambda_m(u-u^m)\|_E^2\\&=\|u-u^m\|_E^2+\frac{2}{\lambda_m}(\Lambda_E^{-1}B^*(y-y^m),u-u^m)_{E^+}\\&+\frac{1}{\lambda_m^2}\|\Lambda_E^{-1}B^*(y-y^m)\|_E^2.\end{aligned} \tag{3.9}$$

On the other hand, substracting (3.4) from (3.2) and multiplying by $y-y^m$ we obtain:

$$\begin{aligned}(\Lambda_E^{-1}B^*(y-y^m),u-u^m)_E&\leqslant-\alpha\|y-y^m\|_V^2-\omega\|\Lambda_E^{-1}B^*(y-y^m)\|_E^2\\&\leqslant-\left(\frac{\alpha}{\|B^*\|^2}+\omega\right)\|\Lambda_E^{-1}B^*(y-y^m)\|_E^2.\end{aligned} \tag{3.10}$$

By using (3.10) in (3.9) it follows

$$\begin{aligned}&\frac{1-2\lambda_m(\omega-\beta)}{\lambda_m^2}\|J^{\omega}_{\lambda_m}(\Lambda_E^{-1}B^*y+\lambda_m u)-J^{\omega}_{\lambda_m}(\Lambda_E^{-1}B^*y^m+\lambda_m u^m)\|_E^2\\&+\left(\frac{1}{\lambda_m^2}-\frac{2\alpha}{\lambda_m\|B^*\|^2}-\frac{2\omega}{\lambda_m}\right)\|\Lambda_E^{-1}B^*(y-y^m)\|_E^2.\end{aligned} \tag{3.11}$$

But by (3.7), there exists $\delta>0$ such that:

$$-\frac{1}{\lambda_m^2}+2\left(\frac{\alpha}{\|B^*\|^2}+\omega\right)\frac{1}{\lambda_m}\geqslant\delta. \tag{3.12}$$

Therefore (3.11) implies the convergence of the sequence $\{\|u-u^m\|_E^2\}$ and then both results imply that:

$$\lim_{m\to\infty}\|\Lambda_E^{-1}B^*(y-y^m)\|_E=0. \tag{3.13}$$

Finally, if $B\Lambda_E^{-1}B^*$ is an isomorphism then (3.8) follows from (3.13), and alternately, if $\alpha>0$, this result holds by (3.10).

Q.E.D.

COROLLARY 3.1

Let $\{\rho_m\}$ be a sequence of real numbers in (0, 1]. We consider the sequence $\{y^m\}$ defined by the following algorithm: u^0 given in E arbitrarily,

$$Ay^m+\omega B\Lambda_E^{-1}B^*y^m=f-Bu^m \tag{3.14}$$

$$u^{m+1}=\rho_m G^{\omega}_{\lambda_m}(\Lambda_E^{-1}B^*y^m+\lambda_m u^m)+(1-\rho_m)u^m. \tag{3.15}$$

Then, under the hypotheses (3.6) and (3.7), we have:

$$\lim_{m\to\infty}\{y^m\} = y. \tag{3.16}$$

Proof

Let $u^{m+(1/2)}$ be defined by:

$$u^{m+(1/2)} = G^{\omega}_{\lambda_m}(\Lambda_E^{-1}B^*y^m + \lambda_m u^m) \tag{3.17}$$

then:

$$\begin{aligned}\|u-u^{m+1}\|_E^2 &= \rho_m^2\|u-u^{m+(1/2)}\|_E^2 + (1-\rho_m)^2\|u-u^m\|_E^2 \\ &\quad + 2\rho_m(1-\rho_m)(u-u^{m+(1/2)}, u-u^m)_E. \end{aligned}\tag{3.18}$$

Similarly to the Proposition 3.1, we obtain

$$\delta\|\Lambda_E^{-1}B^*(y-y^m)\|_E^2 + \|u-u^{m+(1/2)}\|_E^2 \leq \|u-u^m\|_E^2 \tag{3.19}$$

and therefore (3.18) implies:

$$\rho_m^2\delta\|\Lambda_E^{-1}B^*(y-y^m)\|_E^2 + \|u-u^{m+1}\|_E^2 \leq \|u-u^m\|_E^2. \tag{3.20}$$

Now, (3.16) may be obtained as in the last proposition.

COROLLARY 3.2

Under the assumptions of Proposition 3.1 and furthermore, if $\lambda_m = \lambda$ and $\lambda(\omega-\beta) < 1/2$, we have:

$$\lim_{m\to\infty}\{u^m\} = u \quad \text{in } \hat{E} \text{ weakly} \tag{3.21}$$

where

$$u \in \partial\varphi(\Lambda_E^{-1}B^*y).$$

Proof

If $\lambda_m(\omega-\beta) < 1/2$, (3.11) implies:

$$\lim_{m\to\infty}\{J_\lambda^{\omega}(\Lambda_E^{-1}B^*y^m + \lambda u^m)\} = J_\lambda^{\omega}(\Lambda_E^{-1}B^*y + \lambda u) = \Lambda_E^{-1}B^*y. \tag{3.22}$$

Moreover:

$$u^{m+1} - u^m = \frac{\Lambda_E^{-1}B^*y - J_\lambda^{\omega}(\Lambda_E^{-1}B^*y^m + \lambda u^m)}{\lambda} \tag{3.23}$$

hence we have:

$$\lim_{m\to\infty}\{u^{m+1} - u^m\} = 0. \tag{3.24}$$

Then (3.21) holds since the application:

$$v \in E \to G_\lambda^{\omega}(\Lambda_E^{-1}B^*y(v) + \lambda v), \tag{3.25}$$

being $y(v)$ the solution of

$$Ay + \omega B\Lambda_E^{-1}B^*y = f - Bv, \tag{3.26}$$

is non expansive (see Pazy[11] Corollary 4, p. 199).

Q.E.D.

Remark 3.2. For $\omega = (1/2\lambda)$, $E = V$ and $B = \Lambda_V$ the algorithm 1 is equivalent to the algorithm *I* in Lions–Mercier[12].

3.2 *Application to the examples*

Example 2.3.1. In this case it is easy to see that:

$$G_\lambda{}^\omega(v) = \frac{1}{1-\lambda\omega} P_K\left(\frac{1}{\lambda} v\right) - \frac{\omega}{1-\lambda\omega} v, \tag{3.27}$$

where P_K denotes the orthogonal projection on the closed convex K. Consequently, the algorithm 1 is: u^0 given in $(L^2(\Omega))^n$ arbitrary,

$$\begin{aligned} &-(\nu+\omega)\Delta y^m = f + \operatorname{div} u^m \\ &u^{m+1} = \frac{1}{1-\lambda_m\omega} P_K\left(\frac{1}{\lambda_m}\operatorname{grad} y^m + u^m\right) - \frac{1}{1-\lambda_m\omega}(\operatorname{grad} y^m + \lambda_m u^m). \end{aligned} \tag{3.28}$$

By taking $\omega = 0$ in (3.28) we obtain the Uzawa's algorithm for the standard Lagrangian:

$$L(z, v) = \frac{\nu}{2}\int_\Omega |\operatorname{grad} z|^2 \,dx - \int_\Omega fz \,dx + \int_\Omega (v, \operatorname{grad} z)\,dx. \tag{3.29}$$

Example 2.3.2. Now the Yosida approximation of G^ω is:

$$G_\lambda{}^\omega(v) = \frac{1}{\lambda}\left(v - P_K\left(\frac{1}{1-\lambda\omega} v\right)\right) \tag{3.30}$$

and hence the algorithm 1 is:

$$\begin{aligned} &-(\nu+\omega)\Delta y^m = f + \operatorname{div} u^m \\ &u^{m+1} = u^m + \frac{1}{\lambda_m}\operatorname{grad} y^m - \frac{1}{\lambda_m} P_K\left(\frac{1}{1-\lambda_m\omega}(\operatorname{grad} y^m + \lambda_m u^m)\right). \end{aligned} \tag{3.31}$$

3.3 *The case* $\alpha = \beta = 0$

If $\alpha = \beta = 0$, $B\Lambda_E{}^{-1}B^*$ is an isomorphism, $\lambda_m = \lambda$ and $\omega = (1/2\lambda)$, it is also possible to prove the non-expansivity of the application (3.25), by using the techniques of the proposition 3.1. Therefore, the theorem 1 of [13] may be used and hence, if $\{k_m\}$ is a sequence of real numbers such that:

$$\text{(i)}\quad k_m \in [0, 1),\quad \{k_m\} \to 1$$

$$\text{(ii)}\quad \sum_{m\geq 1}(1-k_m) = +\infty$$

$$\text{(iii)}\quad \frac{k_m - k_{m-1}}{(1-k_m)^2} \to 0 \quad \text{as } m \to \infty$$

(for example $k_m = 1-(1/m^\alpha)$, if $0<\alpha<1$), it follows that the sequence $\{u^m\}$ defined by: u^0 arbitrary in E,

$$\begin{aligned} &Ay^m + \frac{1}{2\lambda} B\Lambda_E{}^{-1}B^* y^m = f - Bu^m \\ &u^{m+1} = k_m G_\lambda{}^\omega(\Lambda_E{}^{-1}B^* y^m + \lambda u^m) + (1-k_m)w, \quad w \text{ arbitrary in } E \end{aligned} \tag{3.32}$$

converges strongly to $u \in E$ which satisfies

$$Ay + \frac{1}{2\lambda} B\Lambda_E^{-1}B^*y = f - Bu$$
$$u = G_\lambda^\omega(\Lambda_E^{-1}B^*y + \lambda u) \tag{3.33}$$

being y the solution of (2.10).

From this result it is easy to prove that the sequence $\{y_m\}$ converges to y, since $B\Lambda_E^{-1}B^*$ is an isomorphism.

We give an application of (3.32): Following [5] let T_h be a triangulation of Ω and V_h the subspace of continuous functions in Ω:

$$V_h = \{z_h \in C(\bar{\Omega}):\ z_h|_\Gamma = 0,\ \bar{z}_{h\,K} \in P_1 \quad \text{for all } k \in T_h\}. \tag{3.34}$$

We wish to find $y_h \in V_h$ such that:

$$\int_\Omega (|\text{grad } y_h|^{p-2}\, \text{grad } y_h \cdot \text{grad } z_h)\, dx = \int_\Omega f z_h\, dx \quad \text{for all } z_h \in V_h \tag{3.35}$$

which is an approximation of:

$$-\text{div}\,(|\text{grad } y|^{p-2}\, \text{grad } y) = f \quad \text{in } \Omega$$
$$y = 0 \qquad\qquad \text{on } \Gamma. \tag{3.36}$$

The problem (3.35) is a particular case of (2.10) for the choice:

$$V = V_h,\quad A = 0, E = \left\{v_h : v_h = \sum_{K \in T_h} v_K \chi_K,\ v_K \in R^n\right\}$$
$$B = -\text{div} \quad \text{and} \quad \varphi(v_h) = \frac{1}{p}\int_\Omega |v_h|^p\, dx.$$

Thus, the algorithm (3.32) is:

$$u_h^0 \in E \quad \text{arbitrary,}$$
$$\frac{1}{2\lambda}\int_\Omega (\text{grad } y_h^m \cdot \text{grad } z_h)\, dx = \int_\Omega f z_h\, dx + \int_\Omega \text{div } u_h^m z_h\, dx$$
$$u_h^{m+1} = k_m\left(u_h^m + \frac{1}{\lambda}\,\text{grad } y_h^m - \frac{1}{\lambda} J_\lambda^\omega\,(\text{grad } y_h^m + \lambda_h^m)\right) + (1 - k_m) w_h. \tag{3.37}$$

3.4 *Penalty-duality algorithm*

In the rest of this section we take $\omega = 0$.

Observe that the Lemma 2.1 permits to affirm that y is the solution of the problem (2.10) if and only if there exists $u \in E$ such that:

$$Ay + BG_\lambda(\Lambda_E^{-1}B^*y + \lambda u) = f \tag{3.38}$$
$$u = G_\lambda(\Lambda_E^{-1}B^*y + \lambda u). \tag{3.39}$$

Then, a possible algorithm would be:

Algorithm 2. u^0 is chosen arbitrarily in E,

$$Ay^{m+1} + BG_{\lambda_m}(\Lambda_E^{-1}B^*y^{m+1} + \lambda_m u^m) = f \tag{3.40}$$

$$u^{m+1} = G_{\lambda_m}(\Lambda^{-1}B^*y^{m+1} + \lambda_m u^m). \tag{3.41}$$

PROPOSITION 3.2

If $\{\lambda_m\}$ is an upper bounded sequence of positive real numbers, then the sequence $\{y^m\}$ defined by (3.40) (3.41) converges to the solution of (2.10). Actually, we have:

$$\lim_{m\to\infty} \frac{\|y-y^{m+1}\|_V^2}{\lambda_m} = 0. \tag{3.42}$$

Proof

First observe that:

$$\begin{aligned}\|u-u^{m+1}\|_E^2 = \frac{1}{\lambda_m^2}\|\Lambda_E^{-1}B^*(y-y^{m+1})\|_E^2 + \Big\| u - u^m - \frac{1}{\lambda_m}(J_{\lambda_m}(\Lambda_E^{-1}B^*y+\lambda_m u)\\ - J_{\lambda_m}(\Lambda_E^{-1}B^*y^{m+1}+\lambda_m u^m))\Big\|_E^2 + \frac{2}{\lambda_m}\Big(u-u^{m+1}-\frac{1}{\lambda_m}(J_{\lambda_m}(\Lambda_E^{-1}B^*y+\lambda_m u)\\ - J_{\lambda_m}(\Lambda_E^{-1}B^*y^{m+1}+\lambda_m u^m)), \Lambda_E^{-1}B^*(y-y^{m+1})\Big)_E.\end{aligned} \tag{3.43}$$

On the other hand, from (3.38) and (3.40), it is not difficult to obtain:

$$\begin{aligned}B(u-u^{m+1}) + \frac{1}{\lambda_m} B(J_{\lambda_m}(\Lambda_E^{-1}B^*y+\lambda_m u) - J_{\lambda_m}(\Lambda_E^{-1}B^*y^{m+1}+\lambda_m u^m))\\ = -A(y-y^{m+1}) - \frac{1}{\lambda_m} B\Lambda_E^{-1}B^*(y-y^{m+1}).\end{aligned} \tag{3.44}$$

Consequently (3.43) implies:

$$\begin{aligned}\|u-u^{m+1}\|_E^2 \leqslant -\frac{1}{\lambda_m^2}\|\Lambda_E^{-1}B^*(y-y^{m+1})\|_E^2 - \frac{2\alpha}{\lambda_m}\|y-y^{m+1}\|_V^2\\ + \Big\| u-u^m-\frac{1}{\lambda_m}(J_{\lambda_m}(\Lambda_E^{-1}B^*y+\lambda_m u) - J_{\lambda_m}(\Lambda_E^{-1}B^*y^m+\lambda_m u^m))\Big\|_E^2\end{aligned} \tag{3.45}$$

and by using the Lemma 2.2 for the choices:

$$\omega=-\beta,\quad v_1=u,\quad v_2=u^m,\quad w_1=\Lambda_E^{-1}B^*y+\lambda_m u,\quad w_2=\Lambda_E^{-1}B^*y^m+\lambda_m u^m$$

we get, finally:

$$\begin{aligned}\frac{2\beta}{\lambda_m}\|J_{\lambda_m}(\Lambda_E^{-1}B^*y+\lambda_m u) - J_{\lambda_m}(\Lambda_E^{-1}B^*y^{m+1}+\lambda_m u^m)\|_E^2 + \|u-u^{m+1}\|_E^2\\ + \frac{2\alpha}{\lambda_m}\|y-y^{m+1}\|_V^2 \leqslant \|u-u^m\|_E^2.\end{aligned} \tag{3.46}$$

If $\alpha>0$ this inequality implies (3.42). On the contrary, if $\alpha=0$ but $\beta>0$, then from (3.46) it follows:

$$\lim_{m\to\infty} \frac{\|\Lambda_E^{-1}B^*y - J_{\lambda_m}(\Lambda_E^{-1}B^*y^{m+1}+\lambda_m u^m)\|_E^2}{\lambda_m} = 0 \tag{3.47}$$

and the convergence of the sequence $\{\|u-u^m\|\}$ as $m\to\infty$.

With these results it is easy to prove that

$$\lim_{m\to\infty} \frac{\|\Lambda_E^{-1}B^*(y-y^{m+1})\|_E^2}{\lambda_m} = 0, \tag{3.48}$$

by using (3.45) and the fact that the sequence $\{\lambda_m\}$ is bounded. But, since $B\Lambda_E^{-1}B^*$ is an isomorphism, (3.48) implies (3.42).

Remark 3.3. For $\alpha > 0$, we may give a shorter proof than the previous one for (3.42). Indeed, by using (2.3) we obtain

$$\|u - u^{m+1}\|_E^2 \leq \frac{1}{\lambda_m}(u - u^{m+1}, \Lambda_E^{-1}B^*(y - y^{m+1}) + \lambda_m(u - u^m))_E. \tag{3.49}$$

Moreover, subtracting (3.40) from (3.38) we get:

$$(u - u^{m+1}, \Lambda_E^{-1}B^*(y - y^{m+1}))_E \leq -\alpha\|y - y^{m+1}\|_V^2. \tag{3.50}$$

Substituting this inequality in (3.49) we obtain:

$$\frac{\alpha}{\lambda_m}\|y - y^{m+1}\|_V^2 + \|u - u^{m+1}\|_E^2 \leq \|u - u^m\|_E^2 \tag{3.51}$$

from which it follows (3.42).

Remark 3.4. With the methods of the proposition 3.2, if $\{\rho_m\}$ is a sequence of real numbers in $(0, 2)$, it is not difficult to prove the convergence of the algorithm:

$$\begin{aligned} &Ay^{m+1} + BG_{\lambda_m}(\Lambda_E^{-1}B^*y^{m+1} + \lambda_m u^m) = f \\ &u^{m+1} = \rho_m G_{\lambda_m}(\Lambda_E^{-1}B^*y^{m+1} + \lambda_m u^m) + (1 - \rho_m)u^m. \end{aligned} \tag{3.52}$$

Remark 3.5. If $\lambda_m = \lambda > 0$, we may demonstrate a similar result to the Corollary 3.2, i.e.

$$\lim_{m\to\infty}\{u^m\} = u \quad \text{in } E \text{ weakly} \tag{3.53}$$

where

$$u \in \partial\varphi(\Lambda_E^{-1}B^*y).$$

Remark 3.7. If A is symmetric, the algorithm 2 is exactly the Uzawa's method for the "augmented Lagrangian" (always differentiable):

$$L(z, v) = \frac{1}{2}(Az, z)_{V'V} - (f, z)_{V'V} + \varphi_\lambda(\Lambda_E^{-1}B^*z + \lambda v) - \frac{\lambda}{2}\|v\|_E^2 \tag{3.54}$$

where

$$\varphi_\lambda(v) = \min_{w\in E}\left\{\frac{1}{2\lambda}\|v - w\|_E^2 + \varphi(w)\right\}, \tag{3.55}$$

since

$$\varphi_\lambda' = G_\lambda.$$

Remark 3.8. Observe that the solution of the "regularizated problem":

$$Ay_\lambda + BG_\lambda(\Lambda_E^{-1}B^*y_\lambda) = f \tag{3.56}$$

is exactly the element y^1, obtained by means of the algorithm 2, when $u^0 = 0$ and $\lambda_0 = \lambda$.

Therefore, with the techniques of the Proposition 3.2, we may prove:

$$\|y - y_\lambda\| = 0(\sqrt{\lambda}). \tag{3.57}$$

For example, if φ is the indicator function of the closed convex K, we have $G_\lambda = (I - P_K/\lambda)$ and hence (3.56) is a penalty convergent method when $\lambda \to 0$.

3.5 *Numerical methods for solving the nonlinear problem in the penalty-duality algorithm*

The main difficulty for the practical implementation of the algorithm 2 is the nonlinear problem (3.40).

To solve it, we propose below two iterative methods.

For simplicity we omit the indices. We wish to find the solution of:

$$Ay + BG_\lambda(\Lambda_E^{-1}B^*y + \lambda u) = f \tag{3.58}$$

where now, u is fixed in E.

We may apply the algorithm 1 to this problem; for $\omega = 0$ and $\lambda = \mu$ it is exactly:

$$\begin{aligned} Ay_{p+1} &= f - Bw_p \\ w_{p+1} &= G_{\lambda+\mu}(\Lambda_E^{-1}B^*y_{p+1} + \lambda u + \mu w_p), \quad \mu > 0. \end{aligned} \tag{3.59}$$

If $\mu > (\|B^*\|^2/2\alpha)$, the Proposition 3.1 affirms the convergence of the sequence $\{y_p\}$ to the solution of (3.58); furthermore, in this case, the application which transforms w_p in w_{p+1} is a strict contraction with constant $\sigma = (\mu/\lambda + \mu)$.

Remark 3.9. Since $\mu > (\|B^*\|^2/2\alpha)$, the constant σ goes to one when λ goes to zero, and so the convergence of $\{y_p\}$ may be slow. Thus, it is of interest to take the first values of the sequence $\{\lambda_m\}$ not too small (see Numerical results).

Remark 3.10. Let us consider the algorithm 2. If in solving the problem (3.40) we made only one iteration of the algorithm (3.59) with $w_0 = u^m$, then it is not difficult to show that we obtain exactly the algorithm 1 for $\omega = 0$ and the parameters $\{\lambda_m + \mu\}$.

Now we expose a second algorithm for solving (3.58): y_0 is chosen arbitrarily in V,

$$Ay_{p+1} + \frac{1}{\lambda}B\Lambda_E^{-1}B^*y_{p+1} = f - Bu + \frac{1}{\lambda}J_\lambda(\Lambda_E^{-1}B^*y_p + \lambda u). \tag{3.60}$$

The convergence of this algorithm in established in the following:

PROPOSITION 3.3

If λ is an arbitrary positive real number, we have:

$$\lim_{p\to\infty} \{y_p\} = y \tag{3.61}$$

where y is the solution of (3.58).

Proof

Substracting (3.60) from (3.58) we easily obtain:

$$\alpha\|y - y_{p+1}\|_V^2 + \frac{1}{\lambda}\|\Lambda_E^{-1}B^*(y - y_{p+1})\|_E^2 \leq \frac{1}{\lambda(1+\lambda\beta)}\|\Lambda_E^{-1}B^*(y - y_p)\|_E^2 \tag{3.62}$$

since J_λ is a Lipschitz function with constant $(1/1 + \lambda\beta)$.

From this inequality we deduce immediately (3.61) if $\alpha > 0$.

Alternately, if $\alpha = 0$ but $\beta > 0$, then (3.62) implies:

$$\lim_{p\to\infty} \{\Lambda_E^{-1}B^*y_p\} = \Lambda_E^{-1}B^*y \tag{3.63}$$

from which it follows (3.61), since $B\Lambda_E^{-1}B^*$ is an isomorphism.

Q.E.D.

3.6 *Some variants of the penalty-duality algorithm*

We assume $\lambda_m = \lambda$ (positive constant). If in solving the problem (3.40) through the algorithm

(3.60), we put $y_0 = y^m$ and carry out only one iteration, we obtain the following algorithm:
Let u^0 be chosen arbitrarily in E,

$$Ay^{m+1} + \frac{1}{\lambda} B\Lambda_E^{-1}B^*y^{m+1} = f - Bu^m + \frac{1}{\lambda} BJ_\lambda(\Lambda_E^{-1}B^*y^m + \lambda u^m) \tag{3.64}$$

$$u^{m+1} = u^m + \frac{1}{\lambda}\Lambda_E^{-1}B^*y^{m+1} - \frac{1}{\lambda} J_\lambda(\Lambda_E^{-1}B^*y^{m+1} + \lambda u^m). \tag{3.65}$$

A variant of (3.64) (3.65) would be:
u^0 given arbitrarily in E

$$Ay^{m+1} + \frac{1}{\lambda} B\Lambda_E^{-1}B^*y^{m+1} = f + \frac{1}{\lambda} BJ_\lambda(\Lambda_E^{-1}B^*y^m + \lambda u^m) - Bu^m \tag{3.66}$$

$$u^{m+1} = u^m + \frac{1}{\lambda}\Lambda_E^{-1}B^*y^{m+1} - \frac{1}{\lambda} J_\lambda(\Lambda_E^{-1}B^*y^m + \lambda u^m). \tag{3.67}$$

The numerical results show that the behaviour of both algorithms is analogous. But the last one has the advantage that the term $u^m - (1/\lambda)J_\lambda(\Lambda_E^{-1}B^*y^m + \lambda u^m)$, which is used to solve (3.66), may be employed to compute u^{m+1}.

By making the change of variable:

$$w^m = u^m - \frac{1}{\lambda}\Lambda_E^{-1}B^*y^m \tag{3.68}$$

it is not difficult to check that the algorithm (3.66) (3.67) is, in fact, the algorithm 1, by putting 2λ instead of λ and $\omega = (1/2\lambda)$.

Thus, the proposition 3.1 may be used and we get:

The sequence obtained by means of the algorithm (3.66) (3.67) converges to the solution of the problem (2.10).

Remark 3.11. When $A = 0$, see Gabay–Mercier[14] for another variant of (3.64)(3.65).

3.7 *Application to the examples*

Example 2.3.1. For this example the algorithm 2, is

$$\begin{aligned} &-\nu\Delta y^{m+1} - \operatorname{div} P_K\left(\frac{1}{\lambda_m}\operatorname{grad} y^{m+1} + u^m\right) = f \\ &u^{m+1} = P_K\left(\frac{1}{\lambda_m}\operatorname{grad} y^{m+1} + u^m\right). \end{aligned} \tag{3.69}$$

The convergence of (3.69) is proved in Fortin[6]. On the other hand, the algorithm (3.66) (3.67) is:

$$\begin{aligned} &-\left(\nu + \frac{1}{\lambda}\right)\Delta y^{m+1} = f - \frac{1}{\lambda}\Delta y^m + \operatorname{div} P_K\left(\frac{1}{\lambda}\operatorname{grad} y^m + u^m\right) \\ &u^{m+1} = \frac{1}{\lambda}\operatorname{grad}(y^{m+1} - y^m) + P_K\left(\frac{1}{\lambda}\operatorname{grad} y^m + u^m\right). \end{aligned} \tag{3.70}$$

Example 2.3.2. In this case, the algorithm 2 is

$$\begin{aligned} &-\left(\nu + \frac{1}{\lambda_m}\right)\Delta y^{m+1} + \frac{1}{\lambda_m}\operatorname{div} P_K(\operatorname{grad} y^{m+1} + \lambda_m u^m) = f + \operatorname{div} u^m \\ &u^{m+1} = u^m + \frac{1}{\lambda_m}\operatorname{grad} y^{m+1} - \frac{1}{\lambda_m} P_K(\operatorname{grad} y^{m+1} + \lambda_m u^m) \end{aligned} \tag{3.71}$$

while, the variant (3.66) (3.67) is:

$$-\left(\nu+\frac{1}{\lambda}\right)\Delta y^{m+1}=f+\operatorname{div} u^m-\frac{1}{\lambda}\operatorname{div} P_K(\operatorname{grad} y^m+\lambda u^m)$$
$$u^{m+1}=u^m+\frac{1}{\lambda}\operatorname{grad} y^{m+1}-\frac{1}{\lambda}P_K(\operatorname{grad} y^m+\lambda u^m). \tag{3.72}$$

4. NUMERICAL RESULTS

4.1 *A model problem*

The one dimensional elastoplastic torsion problem (example 2.3.2) has been chosen to compare the different numerical methods exposed in this article, taking:

$$\Omega=(-1,1),\quad \nu=1,\quad f=2. \tag{4.1}$$

In this case, the exact solution is:

$$y(x)=\begin{cases} x+1 & 1\leqslant x<-0.5\\ -x^2+0.75 & -0.5\leqslant x\leqslant 0.5\\ -x+1 & 0.5\leqslant x\leqslant 1.\end{cases} \tag{4.2}$$

In every iteration, the boundary problem has been solved using a classical finite difference method with step $h=(1/10)$.

The convergence test used is:

$$\sum_{i=1}^{20}(y(x_i)-y_i^m)^2<10^{-10}. \tag{4.3}$$

4.2 *Results for algorithm* 1

Table 1 shows the number of iterations needed by algorithm 1 for the different values of ρ, λ and ω.

It is seen that only two iterations are required for the optimal values, which are

$$\lambda=1,\quad \omega=0,\quad \rho=1$$
$$\lambda=0.5,\quad \omega=1,\quad \rho=1$$
$$\lambda=0.75,\quad \omega=0,\quad \rho=0.75.$$

4.3 *Results for algorithm* 2

The number of iterations required by algorithm 2 for the different values of ρ and λ, are exposed in Table 2.

The non-linear problem in (3.71) has been solved with the algorithm (3.60), the convergence test being:

$$\sum_{i=1}^{20}(y_{p+1,i}^{m+1}-y_{p,i}^{m+1})^2<10^{-10}.$$

It is found that for $\rho=1.5$ and $\lambda=0.5$ the number of iterations is two as in algorithm 1.

Even though the results are comparable, it is necessary to remark that every iteration of algorithm 1 needs the solution of a linear problem while each one of algorithm 2 needs the solution of a non-linear one.

Then, it is clear that algorithm 1 is more efficient than algorithm 2, for the chosen problem. Similar results have been obtained for the Bingham fluid problem (example 2.3.1) but they are not shown here.

Table 1.

λ	0.25				0.5				0.75				1				1.25			
ω \ ρ	0.25	0.5	0.75	1	0.5	0.25	0.75	1	0.5	0.25	0.75	1	0.5	0.25	0.75	1	0.5	0.25	0.75	1
0									30	12	2	12	41	18	10	2	53	24	12	9
$\frac{1}{14\lambda}$					35	16	11	40	31	14	8	9	44	19	11	6	55	25	15	10
$\frac{1}{7\lambda}$					39	17	10	19	34	15	8	7	46	21	12	7	58	26	16	10
$\frac{3}{14\lambda}$					41	18	10	12	36	15	8	5	48	22	13	8	60	28	17	11
$\frac{2}{7\lambda}$	60	28	17		41	19	11	8	37	16	9	4	50	23	13	8	62	29	17	12
$\frac{5}{14\lambda}$	62	29	18	72	41	18	10	7	38	17	10	5	51	23	14	9	63	30	18	12
$\frac{3}{7\lambda}$	62	29	18	25	40	18	10	6	40	18	11	6	52	24	14	9	65	30	19	13
$\frac{1}{2\lambda}$	60	28	17	16	41	18	10	2	46	21	12	8	52	24	15	10	65	31	19	13

Table 2.

λ \ ρ	0.25	0.5	0.75	1	1.25	1.5	1.75
0.1	36	16	25	29	28	22	19
0.2	42	19	11	7	7	9	14
0.3	48	22	13	8	4	7	11
0.4		24	15	10	6	5	9
0.5		27	16	11	7	2	7
0.6		29	18	12	8	5	6
0.7		32	20	13	9	6	6
0.8		34	21	15	10	7	4
0.9		36	23	16	11	8	6
1		39	24	17	12	9	7

REFERENCES

1. A. Bermúdez and C. Moreno, Application of pursuit method to optimal control problems. *Appl. Math. Opt.* (To appear).
2. A. Pazy, *Semi-groups of nonlinear contractions in Hilbert space.* In *Problems of Nonlinear Analysis* (Edited by Cremonese). C.I.M.E., Roma (1971).
3. M. R. Hestenes, Multiplier and gradient methods. *J. Opt. Theory Appl.* **4**, 303–320 (1969).
4. M. J. D. Powell, *A method for the nonlinear constraints in minimization problems.* In *Optimization* (Edited by R. Fletcher). Academic Press, New York (1972).
5. R. Glowinski and A. Marroco, *Sur l'approximation par éléments finis d'ordre un, et la résolution par pénalisation-dualité, d'une classe de problèmes de Dirichlet non-linéaires*, pp. 71–76. RAIRO, R2 (1975).
6. M. Fortin, Minimization of some non-differential functionals by the augmented lagrangian method of Hestenes and Powell. *Appl. Math. Opt.* **2**(3), 236–250 (1976).
7. I. Ekeland and R. Temam, *Analyse Convexe et Inéquations Variationnelles.* Dunod, Paris (1974).
8. G. Duvaut and J. L. Lions, *Les Inéquations en Méchanique et en Physique.* Dunod, Paris (1972).
9. R. Glowinski, J. L. Lions and R. Tremolieres, *Analyse Numérique des Inéquations variationnelles.* Dunod, Paris (1976).
10. H. Brezis, Multiplicateur de Lagrange en torsion elastoplastique. *Archiv. Rat. Mech. Anal.* **49**, 32–40 (1972).
11. A. Pazy, On the asymptotic behaviour of iterates of nonexpasive mappings in Hilbert space. *Israel J. Math.* **26**, 197–204 (1977).
12. P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators.* Rapport Interne No. 29, Ecole Polytechnique (1978).
13. P. L. Lions, Approximation de points fixes de contractions. *C.R. Acad. Sci. Paris Ser. A* **284**, 1357–1358 (1977).
14. D. Gabay and B. Mercier, A dual algorithm for the solution of nonlinerar variational problems via finite element approximation. *Comput. Math. Appl.* **2**, 17–40 (1976).